Convergence of the coupled-wave method for metallic lamellar diffraction gratings

Lifeng Li and Charles W. Haggans

Optical Sciences Center, University of Arizona, Tucson, Arizona 85721

Received July 8, 1992; revised manuscript received December 7, 1992; accepted December 8, 1992

Numerical evidence is presented that shows that, for metallic lamellar gratings in TM polarization, the coupled-wave method formulated by Moharam and Gaylord (J. Opt. Soc. Am. A 3, 1780 (1986)) converges slowly. (In some cases, for achieving a relative error of less than 1% in diffraction efficiencies, the number of spatial harmonics retained in the computation must be much greater than 100.) By classification of the modal methods for analyzing diffraction gratings into two distinct categories, the cause for the slow convergence is analyzed and attributed to the use of Fourier expansions to represent the permittivity and the electromagnetic fields in the grating region. The eigenvalues and the eigenfunctions of the modal fields in the grating region, whose accurate determination is crucial to the success of the coupled-wave method, are shown to converge slowly as a result of the use of these Fourier expansions. Despite its versatility and simplicity, the coupled-wave method should be used with caution for metallic surface-relief gratings in TM polarization.

1. INTRODUCTION

Among the many existing rigorous methods for analyzing electromagnetic wave diffraction by gratings, the coupled-wave method (CWM) formulated by Moharam and Gaylord is unique because of its versatility and simplicity. It can be applied to gratings that have continuous permittivity variations (volume gratings) or discontinuous permittivity variations (surface-relief gratings). Its theoretical formulation uses elementary mathematics, and its numerical implementation does not require sophisticated numerical techniques. These features make the CWM popular among many researchers in the field of diffractive optics. Despite its tremendous popularity, however, a comprehensive analysis of the convergence characteristics of the CWM has never been presented. In a paper devoted to a modal method for dielectric gratings with rectangular grooves, Yamakita and Rokushima showed that the CWM converged considerably more slowly than their modal method. For dielectric surface-relief gratings in TE polarization (electric-field vector parallel to the grating grooves), the CWM has been numerically compared with other rigorous methods and has been found to be in good agreement. For metallic surface-relief gratings, however, numerical comparison between the CWM and other rigorous methods is not available in the literature. Experimental verification has been limited to the TE polarization. Therefore a careful convergence study for the use of the CWM for metallic gratings is necessary.

In this paper we analyze the convergence of the CWM for metallic gratings and show that its convergence rate is slow for the TM polarization (magnetic-field vector parallel to the grating grooves), a fact that is unknown to many people who use the method. We reveal that the origin of the slow convergence is the use of Fourier expansions for the permittivity and the electromagnetic field in the grating region. We further demonstrate that these expansions give rise to the slow convergence of the eigenvalues and the eigenfunctions of the modal fields.

2. NUMERICAL EVIDENCE

As evidence of the slow convergence of the CWM for metallic gratings, we consider an arbitrarily chosen case from a paper by Moharam and Gaylord; i.e., we consider two points in their Fig. 3 where the grating groove depth is 1.0 μm. The grating-diffraction configuration is depicted in our Fig. 1 along with the Cartesian coordinate system that we use in this paper. Both the incident (vacuum) wavelength λ and the grating period d are 1.0 μm.
The grating period parameters are incident angle $\theta = 30^\circ$, wavelength $\lambda = 1.0 \, \mu m$, grating period $d = 1.0 \, \mu m$, groove depth $h = 1.0 \, \mu m$, and duty cycle $d_1/d = 0.5$.

The duty cycle of the grating, $d_1/d$, is 0.5. The plane wave is incident upon the highly conducting gold grating at the first-order Littrow (Bragg) angle. From Fig. 3 of Ref. 3, the (negative) first-order and the zeroth-order diffraction efficiencies for TE polarization can be estimated to be about 74% and 13% and those for TM to be about 19% and 73%, respectively. In the following we examine the accuracy of these diffraction efficiencies. Our numerical implementation of the CWM is based on Ref. 6. The truncated matrix $\epsilon_{nm}$ that is involved in the CWM is obtained by numerically inverting $\epsilon_{nm}$, as recommended by Moharam.\footnote{Directly taking the Fourier expansion of $\epsilon^{-1}(x)$ gives inferior numerical results.} The numerical results of the MMME presented in this paper are obtained with a computer program based on Refs. 8–12. Both computer programs have been thoroughly checked against the available results in the literature and against each other.

Figures 2 and 3 show the convergence of the diffraction efficiencies as $N$ increases for TE and TM polarizations, where $N$ denotes the total number of space harmonics retained in the computations. The integer $N$ is henceforth called the truncation order. In our MMME computer program the number of modal fields retained in the computation is set to equal $N$. In these figures, the hollow symbols represent data computed by the CWM, and the filled symbols represent data computed by the MMME. For both polarizations the efficiencies computed with the MMME converge remarkably quickly as $N$ increases. The efficiencies computed with the CWM converge reasonably quickly for the TE polarization (although much more slowly than those computed with the MMME). For the TM polarization, however, the CWM converges very slowly; in fact, it does not begin converging until $N > 40$. In the limit of large $N$, the CWM results converge toward the values obtained with the MMME.

Table 1 lists the numerical values of the TM diffraction efficiencies computed by the CWM and the MMME at truncation orders $N = 65$ and $N = 121$. Assuming that the diffraction efficiencies given by the MMME at $N = 121$ are exact, at $N = 65$ (the truncation order at which Fig. 3 of Ref. 3 was computed) the relative errors in the first- and zeroth-order diffraction efficiencies given by the CWM are 85% and 14%, respectively. Even at the highest truncation order ($N = 121$) the errors are still large, 19% for first-order and 4% for zeroth-order diffraction efficiencies. Keeping in mind that the computation time of the CWM is roughly proportional to $N^3$, one can conclude that the usefulness of the CWM is practically quite limited for highly conducting gratings in TM polarization.

We are grateful to S. S. H. Naqvi and Z. Hatab of the University of New Mexico, who computed, at our request, the TM diffraction efficiencies with their CWM-based computer program and obtained perfect agreement with our CWM data presented in Fig. 3. The poor convergence shown in Fig. 3 is by no means an isolated case. Similar convergence characteristics have been observed with.

Fig. 1. The grating geometry and the diffraction configuration used for studying convergence of the CWM. The values of the parameters are incident angle $\theta = 30^\circ$, wavelength $\lambda = 1.0 \, \mu m$, grating period $d = 1.0 \, \mu m$, groove depth $h = 1.0 \, \mu m$, and duty cycle $d_1/d = 0.5$.

Fig. 2. Convergence of diffraction efficiencies for TE polarization computed by the CWM (hollow symbols) and the MMME (filled symbols).

Fig. 3. Convergence of (a) the negative first-order and (b) the zeroth-order diffraction efficiencies for TM polarization computed by the CWM (hollow symbols) and the MMME (filled symbols).
other metallic gratings that have different angles of incidence, wavelengths, and duty cycles.

### Table 1. Numerical Values of TM Diffraction Efficiencies of the Grating Shown in Fig. 1 Computed at Two Truncation Orders with the CWM and the MMME

<table>
<thead>
<tr>
<th></th>
<th>N = 65</th>
<th></th>
<th>N = 121</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>η_{1} (%)</td>
<td>η_{2} (%)</td>
<td>η_{1} (%)</td>
</tr>
<tr>
<td>CWM</td>
<td>18.815</td>
<td>73.066</td>
<td>12.060</td>
</tr>
<tr>
<td>MMME</td>
<td>10.168</td>
<td>84.836</td>
<td>10.162</td>
</tr>
</tbody>
</table>

### 3. ANALYSIS

#### A. Classification of Modal Methods

In the theory of diffraction gratings, the analytical methods that describe the electromagnetic fields in the grating-groove region by a modal-field expansion are called modal methods. Each modal field in the expansion satisfies two conditions: (1) it is a solution of Maxwell's equations and appropriate boundary conditions, if any, in the grating region $y_m < y < y_M$, where $y_m$ and $y_M$ are the minimum and the maximum ordinates of the surface- or index-modulation region; (2) the modal functions on lines $y = y_m$ and $y = y_M$ form two orthogonal bases in $L^2[0,d]$. Once the modal expansion is obtained, it is matched to the Rayleigh expansions outside the grating regions along the above two lines, and the unknown diffraction-field amplitudes can be solved from the resulting equations.

For the rest of this paper we confine ourselves to the case that the permittivity in the grating region is finite and varies only in the $x$ direction; that is, $\partial \varepsilon/\partial y = 0$. Then it can be shown\(^{6}\) that the modal expansion can be written as

$$F(x,y) = \sum_n (a_n \cos \lambda_n y + b_n \sin \lambda_n y)u_m(x),$$

where $a_n$ and $b_n$ are constant modal-field amplitudes and $\lambda_n$ and $u_m(x)$ are modal eigenvalues and modal eigenfunctions, which are determined by the boundary-value problem

$$Lu_m(x) = \rho_m u_m(x),$$

where $L$ is a different operator whose expression can be found in Ref. 12, $\alpha_0 = 2\pi \sin \theta \lambda$, and

$$\rho_m = \lambda_m^2.$$

The second equation in Eqs. (2) is called the pseudo-periodic condition.

We call the solution method that solves the scalar boundary-value problem [Eqs. (2)] directly an MMME. The MMME has been used by Collin,\(^{13}\) Botten et al.,\(^{5-10}\) Sheng et al.,\(^{14}\) Suratteau et al.,\(^{11}\) Yamakita and Kokushima,\(^{4}\) Peng,\(^{15}\) and Li\(^{12}\) for lamellar gratings and by Tamir et al.,\(^{16}\) and Chu and Kong\(^{17}\) for sinusoidally varying-index volume gratings. The advantages of the MMME are its accuracy and its efficiency. For certain profiles, the eigenvalues and the eigenfunctions can be determined to great accuracy with ease. Because Eqs. (2) are scalar, each eigenvalue is determined independently of the others, and the computational effort for $M$ solutions is proportional to $M$. The limitation of the MMME is its narrow range of applicability. Efficient numerical methods for solving Eqs. (2) are known for only a few forms of the dielectric function $\varepsilon(x)$, namely, piecewise-constant and sinusoidal forms.

An alternative approach for implementing the modal method is outlined as follows. Because $u_m(x)$ is pseudo-periodic, it can be written as a Fourier expansion,

$$u_m(x) = \sum_n c_{nm} \exp(i\alpha_n x),$$

where

$$\alpha_n = \alpha_0 + 2\pi n/d.$$

[Strictly speaking, the expansion on the right-hand side of Eq. (4) is not a Fourier expansion; it is a product of a Fourier expansion and the function $\exp(i\alpha_n x)$.] Substituting Eq. (4) into Eq. (1), interchanging the summation order, and then completing the summation over $m$, we have

$$F(x,y) = \sum_n s_n(y)\exp(i\alpha_n x),$$

where $s_n$ are the unknown, $y$-dependent spatial-harmonics amplitudes. Substitution of Eq. (6) into Maxwell's equations and the use of the Fourier expansions of $\varepsilon(x)$ lead to an infinite-dimensional, second-order matrix differential equation, or two coupled first-order matrix differential equations for $s_n(y)$. Since $\varepsilon$ is independent of $y$, $s_n$ can take the form

$$s_n(y) = c_n \exp(i\lambda y),$$

and the matrix differential equation(s) can be reduced to an algebraic eigenvalue equation. Symbolically, we write

$$A\nu = \rho \nu,$$

or

$$B\xi = \lambda \xi,$$

where $A$ and $B$ are square matrices, $B$ being twice as large as $A$; $\rho$ and $\lambda$ are eigenvalues; and $\nu$ and $\xi$ are column vectors whose components are $c_n$. Eigenvalues $\rho$ and $\lambda$ are related as in Eq. (3). For each eigenvalue of Eq. (8) or Eq. (9), its associated eigenvector corresponds to an eigenfunction $u_m(x)$ in Eq. (4). To see this, one needs only to substitute the components of the eigenvector $\nu_m$, by using Eq. (7), into Eq. (6) and to compare the result with Eq. (4). We call the solution method described above an MMFE. The MMFE has been used by Burckhardt\(^{18}\) for volume gratings and later by Knop\(^{19}\) for lamellar gratings. In both cases, eigenvalue equations of the type of Eq. (8) were used. The CWM is an MMFE that uses Eq. (9). In our opinion, from a numerical point of view it is a disadvantage to use Eq. (9), because it increases computational effort and computer memory in assembling a larger matrix, and Eq. (9) does not take advantage of the double degeneracy of the eigenvalues as indicated by Eq. (3), unless special care is taken during numerical implementation.

Clearly the MMFE and the MMME are mathematically equivalent. Numerically, however, they are quite differ-
L. Li and C. W. Haggans

Consequences of Using Fourier Expansions

The CWM, or the MMFE in general, expands both the permittivity and the electromagnetic fields in the grating region in Fourier series. The most obvious consequence of using these expansions is that these expansions change the physical problem to be solved, as schematically illustrated in Fig. 4. The original physical structure to be treated is a surface-relief grating that has discontinuous permittivity $\varepsilon(x)$. As a result of the use of the truncated Fourier expansion for $\varepsilon(x)$, the lamellar grating becomes a grating that has continuously varying $\varepsilon(x)$. The MMFE introduces two convergence processes: that of the permittivity and that of the modal fields. For a finite $N$, none of the modes satisfies Maxwell's equations and the original boundary conditions exactly. As the truncation order $N$ increases, the permittivity distribution approaches the original surface-relief profile, and, at the same time, the eigenvalues and the eigenfunctions of the modal fields approach their true values. In contrast, in the MMME, the lamellar-grating geometry is not changed and the modal fields are determined individually and exactly.

Convergence of the Fourier Series

The analysis in Subsection 3.B implies that the convergence rate of the CWM depends on the convergence rate of the Fourier series. For a lamellar grating the permittivity function $\varepsilon(x)$ is given by

$$\varepsilon(x) = \begin{cases} \varepsilon_1 & |x| < d_1/2, \\ \varepsilon_2 & d_1/2 < |x| < d/2. \end{cases}$$  (10)

It is elementary to show that the asymptotic form of the Fourier expansion coefficient, $\varepsilon_n$, of $\varepsilon(x)$ is

$$\varepsilon_n = O(1/n).$$  (11)

Let $u(x)$ be a modal eigenfunction of the $z$ (parallel to the grating grooves) component of the electric field in the TE diffraction problem or a model eigenfunction of the $z$ component of the magnetic field in the TM diffraction problem. One can obtain the asymptotic form of the Fourier coefficient $U_n$ of $u(x)$ from the definition of the Fourier coefficient,

$$U_n = \int_{-d/2}^{d/2} u(x)e^{-i\alpha_n x}dx,$$

by using integration by parts twice and then using Eqs. (2). The results are

$$U_n = O(1/n^3), \quad \text{TE},$$  (12)

$$U_n = O(1/n^2), \quad \text{TM}.$$  (13)

The convergence rates given in Eqs. (11–13) are slow and are responsible for the poor convergence of the CWM.

The difference between Eqs. (12) and (13) is in agreement with the fact that the CWM converges faster for TE than for TM polarization. The difference in convergence rates is due to the difference in the interface conditions for the two polarizations at the discontinuities of $\varepsilon(x)$. For TE polarization both the $z$ components of the electric field and its derivative are continuous. For TM polarization the $z$ component of the magnetic field is continuous, but its derivative is not. Instead we have

$$\left[ \frac{1}{\varepsilon} \frac{du}{dx} \right]_+ = \left[ \frac{1}{\varepsilon} \frac{du}{dx} \right]_-, \quad \text{TM},$$  (14)

where $+$ and $-$ indicate the limits taken from the left-hand side and the right-hand side, respectively, of a discontinuity of $\varepsilon(x)$. By using Eq. (14) we can further rationalize the convergence difference of the CWM for dielectric and metallic gratings in TM polarization. Suppose that $\varepsilon = 1$. For a lossless dielectric grating, $\varepsilon > 0$. So although the derivative of the modal field is discontinuous, the two limiting values of the derivatives have the same sign. For a metallic grating, however, $\varepsilon < 0$, if the small imaginary part of $\varepsilon$ is neglected. Accordingly, the two limiting values of the derivatives have opposite signs. Thus, at the permittivity discontinuity, the TM modal fields are said to be smoother in a dielectric grating than they are in a metallic grating. In other words, the TM modal fields have smaller higher-order Fourier coefficients in a dielectric grating than they do in a metallic grating. Hence in TM polarization the CWM converges faster for dielectric gratings than it does for metallic gratings.

Note that the permittivity expansion is used for both TE and TM polarizations, and its convergence rate is slower than the convergence rates of the field expansions for both polarizations. This fact poses a difficulty in understand-
expected that the convergence of the eigenvalues will be

Fig. 5. Convergence of the tenth eigenvalue, −ρ0, for TE polarization. Circles represent the real part, triangles represent the imaginary part, and dashed lines mark the exact values.

Fig. 6. Convergence of the ninth eigenvalue, −ρ0, for TM polarization. Circles represent the real part, triangles represent the imaginary part, and dashed lines mark the exact values.

ing why the MMFE converges more slowly for TM than for TE, because the convergence rate of the permittivity expansion might be the dominant factor in determining the convergence rate of the MMFE. However, we notice that the permittivity expansion appears in Eq. (8) or Eq. (9) quite differently for TE versus TM polarizations. In addition, and perhaps more important, we cannot assume a priori that the convergence rate of the permittivity expansion and that of the (TE or TM) field expansion have equal influence on the convergence rate of the diffraction efficiencies. We admit that the arguments given in this subsection regarding the TE and the TM convergence-rate difference are a posteriori. The convergence-rate difference cannot be completely explained by such a simplistic convergence analysis of the Fourier expansions. One may have to look into the field structural difference in the grating region for the two polarizations. As a final note in this subsection, we mention that the TE and the TM convergence-rate difference also exists, although less prominently, with the integral method, which does not use Fourier expansions for the permittivity and the field.

D. Convergence of the Eigensolutions
In the CWM the eigenvalues are solved from the matrix eigenvalue equation, which is obtained from the truncated and the slowly converging Fourier expansions of the permittivity and the unknown modal fields. Thus it can be expected that the convergence of the eigenvalues will be slow. Figures 5 and 6 show the convergence of the real and the imaginary parts, respectively, of −ρ0 for TE polarization and −ρ0 for TM polarization. (The selection of these two eigenvalues was arbitrary. The numbering of the modal eigenvalues is in ascending order in modulus, and it has no direct relationship to the numbering of diffraction orders outside the grating grooves.) The grating geometry and the physical parameters are those of Fig. 1. Similarly to the convergence of diffraction efficiencies shown in Figs. 2 and 3, the TE eigenvalue computed with the CWM converges reasonably quickly toward that computed with the MMME, but the TM eigenvalue converges slowly. During our numerical experimentation, we noticed that all the TE eigenvalues computed with the CWM converge at about the same rate as the eigenvalue displayed in Fig. 5, and about half of the TM eigenvalues converge faster than the one displayed in Fig. 6.

We do not include figures showing the convergence of the eigenfunctions, because it is difficult to present graphically the convergence of a sequence of functions. Nevertheless, since the numerical value of an eigenfunction depends keenly on its associated eigenvalue, the convergence behavior of the eigenfunctions can be expected to be the same as that of the eigenvalues.

In general, the difficulty in using any method in grating analysis lies in the accurate and efficient determination of the electromagnetic field in the grating region. For the modal methods the total field is given by the superposition of the modal fields, and the latter are determined from their associated eigenvalues. Therefore the accuracy and the convergence rate of the overall grating program depend on the accuracy and the convergence rate of the eigenvalues. The poor convergence of the TM eigenvalues displayed in Fig. 6 explains the poor convergence of diffraction efficiencies displayed in Fig. 3.

4. CONCLUSIONS
In this paper we have provided strong numerical evidence showing that the coupled-wave method converges slowly for metallic gratings in TM polarization. Our analysis shows that the slow convergence is caused by the use of the slowly convergent Fourier expansions for the permittivity and the electromagnetic fields in the grating region. We classified the modal methods into two categories, MMME and MMFE. This classification facilitated our analysis of the consequences of using the Fourier expansions. We also studied the convergence of the eigenvalues computed by the CWM. We conclude from this analysis that despite its versatility and simplicity, the coupled-wave method should be used with caution for metallic surface-relief gratings in TM polarization.

For the sake of simplicity, we dealt only with lamellar gratings in this paper. However, our analysis and conclusions apply to the use of the CWM for surface-relief gratings in general, because an arbitrary-profile surface-relief grating can be approximated by a stack of lamellar gratings.

ACKNOWLEDGMENTS
This research was supported by the Optical Data Storage Center at the University of Arizona. The research of
C. W. Haggans was also supported by a graduate fellowship from Itek Optical Systems.

*Present address, 3M Company, 1185 Wolters Boulevard, Vadnais Heights, Minnesota 55110.

REFERENCES